Theoretical Computer Science 560 (2014) 91–106

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Optimal axis compensation in quantum key distribution protocols over unital channels ☆

Shun Watanabe [a,b,*], Ryutaroh Matsumoto [c], Tomohiko Uyematsu [c]

[a] Department of Information Science and Intelligent Systems, University of Tokushima, 2-1, Minami-josanjima, Tokushima, 770-8506, Japan
[b] Institute for Systems Research, University of Maryland, College Park, MD 20742, USA
[c] Department of Communications and Computer Engineering, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo, 152-8552, Japan

ARTICLE INFO

Article history:
Received 24 February 2009
Accepted 8 July 2013
Available online 19 September 2014

Keywords:
Axis compensation
BB84 protocol
Quantum key distribution
Unital channel

ABSTRACT

The axis compensation is a procedure in which the sender and the receiver compensate the axes of their transmitter and detector so that the bit sequence can be transmitted more reliably. We show the optimal axis compensations maximizing the key generation rate for unital channels. We consider the case in which only Bob is allowed to compensate his axis, and the case in which both Alice and Bob are allowed to compensate their axes. In the former case, we show that we should utilize the mismatched measurement outcomes in the channel estimation phase. In the latter case, we show that we do not have to utilize the mismatched measurement outcomes in the channel estimation phase.

1. Introduction

Quantum key distribution (QKD) has attracted great attention as a technology to realize the information theoretically secure key agreement. In this paper, we investigate the Bennett–Brassard 1984 (BB84) protocol [1] and the six-state protocol [5], and the QKD protocols indicate the BB84 protocol and the six-state protocol.

Typically in theoretical studies on the QKD protocols, the protocols roughly consist of three phases: the bit transmission phase, the channel estimation phase, and the postprocessing phase. In the bit transmission phase, the legitimate sender, usually referred to as Alice, sends a bit sequence to the legitimate receiver, usually referred to as Bob, by encoding them into qubits. After this phase, Alice and Bob publicly announce the bases they used for each qubit. In this paper, those bits that are transmitted and received by the same bases are called the matched measurement outcomes, and those bits that are transmitted and received by different bases are called the mismatched measurement outcomes. In the standard QKD protocols, the mismatched measurement outcomes are discarded, and the matched measurement outcomes are called the sifted key. In the channel estimation phase, Alice and Bob estimate the channel and the amount of information gained by an eavesdropper, usually referred to as Eve. For the channel estimation, we sacrifice a part of the sifted key, which is usually called the test bits. Finally in the postprocessing phase, Alice and Bob share a secret key based on their bit sequences

☆ A part of this paper was presented at 2009 IEEE International Symposium on Information Theory held in Seoul, Korea, from June 28 through July 3, 2009, and was published in pages 1884–1888 of the proceedings. The results in Section 3.2.2 are additionally shown in this paper.

* Corresponding author at: Department of Information Science and Intelligent Systems, University of Tokushima, 2-1, Minami-josanjima, Tokushima, 770-8506, Japan.
E-mail addresses: shun-wata@is.tokushima-u.ac.jp (S. Watanabe), uyematsu@ieee.org (T. Uyematsu).
URL: http://www.rmatsumoto.org/research.html (R. Matsumoto).

http://dx.doi.org/10.1016/j.tcs.2014.09.020
0304-3975/

Table 1
Summary of the optimized key generation rates for various settings.

Channel estimation			Accurate	Conventional
Six-state	One-side		$F_1(\mathcal{E})$	$\tilde{F}_1(\mathcal{E})$
	Two-side		$F_2(\mathcal{E})$	$\tilde{F}_2(\mathcal{E})$
BB84	z–x plane	One-side	$G_1(\mathcal{E})$	$\tilde{G}_1(\mathcal{E})$
		Two-side	$G_2(\mathcal{E})$	$\tilde{G}_2(\mathcal{E})$
	Any direction	One-side	$J_1(\mathcal{E})$	$\tilde{J}_1(\mathcal{E})$
		Two-side	$J_2(\mathcal{E})$	$\tilde{J}_2(\mathcal{E})$

obtained in the bit transmission phase. Usually, the postprocessing phase consists of two procedures: the information reconciliation, which is also known as the error correction, and the privacy amplification. In the following, we sometimes call these three phases the key agreement phase as a whole.

On the other hand, in the practical QKD protocols, Alice and Bob conduct the *axis compensation* (before the bit transmission phase), in which Alice and Bob compensate the axes of their transmitter and detector so that the bit sequence can be transmitted more reliably in the bit transmission phase. This axis compensation is considered to be indispensable in the QKD protocols, and it has been extensively studied from the experimental point of view [7,10,18,28,29,31,24,30,6,16] (see also [13]). However, it has not been theoretically clarified how Alice and Bob should compensate the axes in the axis compensation.

In this paper, we investigate the optimal axis compensation in the sense that the key generation rate is maximized, where the key generation rate is defined as the ratio between the length of the shared secret key and that of the sequences initially possessed by Alice and Bob in the postprocessing phase.

We consider the following various settings. In the channel estimation phase, we consider two kinds of channel estimation: the accurate channel estimation and the conventional channel estimation (see [26]). In the accurate channel estimation, we use the mismatched measurement outcomes in addition to the matched measurement outcomes to estimate the channel. On the other hand, in the conventional channel estimation, we discard the mismatched measurement outcomes and only use the matched measurement outcomes. The reason why we consider two kind of channel estimation is that the authors recently clarified that the key generation rate is increased if we use the accurate channel estimation instead of the conventional channel estimation [26]. It is worthwhile to clarify whether we should use the accurate channel estimation instead of the conventional channel estimation when the QKD protocols involve the axis compensation.

In the postprocessing phase, we employ the standard postprocessing. We do not use the noisy preprocessing [22,15] nor the two-way classical communication [14,27].

In the axis compensation phase, we consider two kinds of compensations:

(i) (*one-side compensation*) Only Bob is allowed to compensate his axis.
(ii) (*two-side compensation*) Both Alice and Bob are allowed to compensate their axes.

Furthermore in the BB84 protocol, we subdivide each compensation into two kinds. In the first kind, Bob (or both Alice and Bob) is (are) allowed to compensate his (their) axis within the z–x plane of the Bloch sphere. In the second kind, Bob (or both Alice and Bob) is (are) allowed to compensate his (their) axis within any direction. The reason why we consider these two kind of compensations in the BB84 protocol is as follows. Since we only use the z-basis and x-basis in the BB84 protocol, it is natural to consider the axis compensation within the z–x plane. On the other hand, we might use the axis compensation within any direction if we are allowed to enhance the device for the compensation. Indeed, several researchers employ the compensation within any direction in the literature [7,10,18,28,29,31,24,30,6,16]. Therefore, we also investigate the compensation within any direction.

The optimized key generation rates (of the standard postprocessing) $F_1(\mathcal{E})$, $\tilde{F}_1(\mathcal{E})$, $F_2(\mathcal{E})$, $\tilde{F}_2(\mathcal{E})$, $G_1(\mathcal{E})$, $\tilde{G}_1(\mathcal{E})$, $G_2(\mathcal{E})$, $\tilde{G}_2(\mathcal{E})$, $J_1(\mathcal{E})$, $\tilde{J}_1(\mathcal{E})$, $J_2(\mathcal{E})$, $\tilde{J}_2(\mathcal{E})$ for above described 12 settings are summarized in Table 1. These quantities are formally defined in Sections 2.2 and 2.3 respectively.

In this paper, we investigate the above described optimized key generation rates, and derive closed-form expression of $F_1(\mathcal{E})$, $F_2(\mathcal{E})$, $G_1(\mathcal{E})$, $G_2(\mathcal{E})$, $J_1(\mathcal{E})$, and $J_2(\mathcal{E})$ for unital channels. The reason why we concentrate on the unital channel is as follows. The Pauli channel is one of the most important class of channels in the study of the QKD protocols and is well investigated in the literatures. When the channel is a Pauli channel and the axes of Alice's transmitter and Bob's detector are misaligned, then the resulting channel can be regarded as a unital channel. Thus the unital channel deserves consideration in the QKD research as well as the Pauli channel. There is also a technical reason why we deal with unital channels. For unital channels, we have closed-form expressions of key generation rates of the QKD protocols. The existence of such closed-form expressions enables us to identify optimal compensation procedures. Without them identification is difficult. Although the main results in this paper are stated under the assumption that the channel is unital, it should be noted that our protocol is secure for any channel (see Remark 5).

There are also other assumptions that should be noted. In this paper, we assume that the channel is constant over the time, i.e., the so-called collective attack, and we also assume that the channel is unchanged in the compensation phase and the key agreement phase. However, our protocol is secure even if these assumptions are not satisfied, because the key

agreement phase itself is just an existing secure QKD protocol[1] [21,26]. If the channel changes arbitrarily, we cannot discuss the optimality of the compensation procedures. Thus, these assumptions are needed for our problem to be meaningful.

By using the closed-form expressions of the optimized key generation rates, we also derive the following relationships:

$$F_2(\mathcal{E}) = \tilde{F}_2(\mathcal{E}) = F_1(\mathcal{E}) \geq \tilde{F}_1(\mathcal{E}),$$
$$G_2(\mathcal{E}) = \tilde{G}_2(\mathcal{E}) = G_1(\mathcal{E}) \geq \tilde{G}_1(\mathcal{E}),$$
$$J_2(\mathcal{E}) = \tilde{J}_2(\mathcal{E}) \geq J_1(\mathcal{E}) \geq \tilde{J}_1(\mathcal{E})$$

hold for any unital channel, and

$$F_1(\mathcal{E}) > \tilde{F}_1(\mathcal{E}),$$
$$G_1(\mathcal{E}) > \tilde{G}_1(\mathcal{E}),$$
$$J_1(\mathcal{E}) > \tilde{J}_1(\mathcal{E})$$

hold for general cases of unital channels.

Our results provide the following important insight. In the literatures [7,10,18,28,29,31,24,30,6,16], they employ the one-side compensation for the axis compensation phase and the conventional channel estimation for the channel estimation phase. However, when we employ the one-side compensation, above mentioned relationships imply that we should use the accurate channel estimation. On the other hand, when we employ the two-side compensation, above mentioned relationships imply that we do not have to use the accurate channel estimation.

It should be noted that the axis compensation is not the only method to combat asynchronous axes, but there are some other methods. One method is to encode a bit into a decoherence-free subspace of multiple qubits so that Bob can detect the transmitted qubits reliably without synchronization of the axes [25,4]. This method has a drawback such that encoding a bit into multiple qubits is not practical by the present day technology.

Recently, another method that does not use the axis compensation was also proposed in [17]. In this method, a bit is encoded into a single qubit, and the transmitted qubit is detected without synchronization of the axes. After the agreement of bit sequence, Alice and Bob estimate the channel by using only statistics that are independent of the angle between the axes. Since this methods only uses a part of the statistics, the key generation rate might be smaller than that of the method using the axis compensation.

The two methods that do not use the axis compensation can be used even if the variation of the angle between the axes is fast as in the free space QKD. On the other hand, if the variation of the angle between the axes is slow as in the optical fibre QKD, the method using the axis compensation seems more eligible than the other two because of the above mentioned reasons.

The rest of this paper is organized as follows: In Section 2, we formally describe the problem mentioned above for the six-state protocol and the BB84 protocol. In Section 3, we provide closed-form expressions of the optimized key generation rates, and also clarify the relationships among the optimized key generation rates for various settings. We state the conclusion in Section 4. Proofs of main theorems are shown in Appendices A–F.

2. Problem formulation

In this section, we formally describe the problem we investigate in this paper. Suppose that Alice and Bob are connected by a qubit channel $\mathcal{E}_B$ from the set of all qubit density operators to themselves. As is usual in QKD literatures, we assume that Eve can access all the environment of channel $\mathcal{E}_B$; the channel to the environment is denoted by $\mathcal{E}_E$. In the rest of this paper, we omit the subscripts B and E if they are obvious from the context.

It should be noted that $\mathcal{E}$ can be any qubit channel throughout the paper, unless we specify the channel to be a Pauli channel or a unital channel.

2.1. Stokes parameterization and Choi operator

For convenience, we introduce the Stokes parameterization and the Choi operator for the qubit channel. The qubit channel $\mathcal{E}$ can be described by the affine map parameterized by 12 real parameters [12,11] as follows:

$$\begin{bmatrix} \theta_z \\ \theta_x \\ \theta_y \end{bmatrix} \mapsto \begin{bmatrix} R_{zz} & R_{zx} & R_{zy} \\ R_{xz} & R_{xx} & R_{xy} \\ R_{yz} & R_{yx} & R_{yy} \end{bmatrix} \begin{bmatrix} \theta_z \\ \theta_x \\ \theta_y \end{bmatrix} + \begin{bmatrix} t_z \\ t_x \\ t_y \end{bmatrix}, \tag{1}$$

[1] In the usual security proof, the security of the QKD protocol is proved under the assumption that all information that can be obtained outside the QKD protocol is known to Eve. Thus, Eve cannot get any advantage even if we conduct the compensation procedure.

where $(\theta_z, \theta_x, \theta_y)$ describes a vector in the Bloch sphere [19]. The pair (R, t) of the matrix and the vector in Eq. (1) is called the Stokes parameterization of the channel $\mathcal{E}$. In the rest of this paper, we identify the channel $\mathcal{E}$ and its Stokes parameterization, and occasionally write $\mathcal{E} = (R, t)$.

For the channel $\mathcal{E}$ and each pair of bases $(a, b) \in \{z, x, y\}^2$, define the biases of the outputs as

$$Q_{ab0} := \langle 0_b|\mathcal{E}_B\big(|0_a\rangle\langle 0_a|\big)|0_b\rangle - \langle 1_b|\mathcal{E}_B\big(|0_a\rangle\langle 0_a|\big)|1_b\rangle,$$
$$Q_{ab1} := \langle 1_b|\mathcal{E}_B\big(|1_a\rangle\langle 1_a|\big)|1_b\rangle - \langle 0_b|\mathcal{E}_B\big(|1_a\rangle\langle 1_a|\big)|0_b\rangle,$$

where $|0_a\rangle$, $|1_a\rangle$ are eigenstate of the Pauli operator σ_a for $a \in \{x, y, z\}$ respectively. Then, a straight forward calculation shows the relations

$$R_{ba} = \frac{1}{2}(Q_{ab0} + Q_{ab1}), \qquad t_b = \frac{1}{2}(Q_{ab0} - Q_{ab1}). \tag{2}$$

A unital channel is a channel that maps the completely mixed state $I/2$ to itself. For a unital channel, the vector part (t_z, t_x, t_y) of the Stokes parameterization is the zero vector. Furthermore, the channel is called a Pauli channel if the matrix part R of the Stokes parameterization is a diagonal matrix in addition to that the vector part is the zero vector.

We can also describe the qubit channel $\mathcal{E}$ by the Choi operator

$$\rho_{AB} := (\mathrm{id} \otimes \mathcal{E}_B)\big(|\psi\rangle\langle\psi|\big),$$

where $|\psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ is a maximally entangled state. For the unital channel, the Choi operator satisfies $\mathrm{Tr}_A[\rho_{AB}] = I/2$. Furthermore, the channel is a Pauli channel if and only if the Choi operator is a Bell diagonal operator, i.e.,

$$\rho_{AB} = \sum_{a\in\{i,z,x,y\}} q_a|\psi_a\rangle\langle\psi_a|$$

for Bell states $|\psi_a\rangle := (I \otimes \sigma_a)|\psi\rangle$ and the probability distribution (q_i, q_z, q_x, q_y) on $\{i, z, x, y\}$, where σ_i is the identity operator. Throughout this paper, we omit the subscript AB if it is obvious from the context.

2.2. Six-state protocol

At the preparation phase of the six-state protocol, Alice and Bob conduct the following axis compensation procedure. Alice randomly sends bit 0 or 1 to Bob by modulating it into a transmission basis that is randomly chosen from the z-basis $\{|0_z\rangle, |1_z\rangle\}$, the x-basis $\{|0_x\rangle, |1_x\rangle\}$, or the y-basis $\{|0_y\rangle, |1_y\rangle\}$. Then Bob randomly chooses one of measurement observables σ_x, σ_y, and σ_z, and converts a measurement result $+1$ or -1 into a bit 0 or 1 respectively. After a sufficient number of transmissions, Alice and Bob publicly announce their transmission bases and measurement observables. They also announce all of their bit sequences for estimating channel $\mathcal{E}$. Note that Alice and Bob do not discard mismatched measurement outcomes, which are transmitted and received by different bases, and they also use the mismatched measurement outcomes to estimate the channel. From Eq. (2), we find that Alice and Bob can estimate all parameters (R, t) of the channel in the six-state protocol [8,20]. Note that the use of the mismatched measurement outcomes for channel estimation is also known as the process tomography.

We consider two kinds of compensations:

(i) Only Bob is allowed to compensate his axis, i.e., the channel after the compensation is

$$\mathcal{E}'_B = \mathcal{U}_B \circ \mathcal{E}_B, \tag{3}$$

where $\mathcal{U}_B$ is a unitary channel that represents Bob's compensation.

(ii) Both Alice and Bob are allowed to compensate their axes, i.e., the channel after the compensation is

$$\mathcal{E}'_B = \mathcal{U}_B \circ \mathcal{E}_B \circ \mathcal{U}_A, \tag{4}$$

where $\mathcal{U}_A$ and $\mathcal{U}_B$ are unitary channels that represent Alice and Bob's compensations.

Based on the estimate of the parameters (R, t), Bob (or both Alice and Bob) decides $\mathcal{U}_B$ (or $\mathcal{U}_A$ and $\mathcal{U}_B$), and he (they) compensates the channel. The choice of $\mathcal{U}_B$ (or $\mathcal{U}_A$ and $\mathcal{U}_B$) can be decided according to Theorem 4 and Corollary 6 of Section 3.1.

Remark 1. Throughout this paper, the prime represents that it is after the compensation. ■

After the compensation procedure, Alice and Bob conduct the above bit transmission and reception procedure again. This time, they only announce a part of their bit sequence for estimating the channel $\mathcal{E}'$, and they conduct the postprocessing to generate a secret key from the remaining (unannounced) bit sequences.

Henceforth, we focus on the postprocessing procedure for Alice's bit sequence $\mathbf{x} \in \mathbb{F}_2^n$ that is transmitted in z-basis and corresponding Bob's bit sequence $\mathbf{y} \in \mathbb{F}_2^n$ that is received in σ_z measurement, where $\mathbb{F}_2$ is the finite field of order 2. We employ the standard postprocessing procedure that consists of the information reconciliation procedure and the privacy amplification procedure (e.g. see [26, Section 2]). Note that we do not use the so-called noisy preprocessing [22,15] nor the postprocessing with two-way classical communication [14,27].

Let

$$H_{\mathcal{E}'}(X|E) := H\big(\rho'_{XE}\big) - H\big(\rho'_E\big)$$

be the conditional von Neumann entropy with respect to the density operator

$$\rho'_{XE} := \sum_{x \in \mathbb{F}_2} \frac{1}{2} |x_z\rangle\langle x_z| \otimes \mathcal{E}'_E\big(|x_z\rangle\langle x_z|\big)$$

on the joint system $\mathcal{H}_X \otimes \mathcal{H}_E$, where $H(\rho)$ is the von Neumann entropy [19] for a density matrix ρ, and we take the base of the logarithm to be 2 throughout the paper. For the compensated channel $\mathcal{E}'$, we define the joint probability distribution

$$P'_{XY}(x, y) := \frac{1}{2} \langle y_z|\mathcal{E}'_B\big(|x_z\rangle\langle x_z|\big)|y_z\rangle$$

of the joint random variable (X, Y) on $\mathbb{F}_2 \times \mathbb{F}_2$. Then, let

$$H_{\mathcal{E}'}(X|Y) := -\sum_{x,y \in \mathbb{F}_2} P'_{XY}(x, y) \log P'_{X|Y}(x|y) \tag{5}$$

be the conditional Shannon entropy of X given Y.

In the six-state protocol, since Alice and Bob can estimate the channel $\mathcal{E}'$ exactly if they use the accurate channel estimation, they can asymptotically share a secure key if the length ℓ of the key satisfies

$$\frac{\ell}{n} < H_{\mathcal{E}'}(X|E) - H_{\mathcal{E}'}(X|Y)$$

(see [21,26]). Therefore, we consider the following two optimization problems:

(i) Find a closed-form expression of

$$F_1(\mathcal{E}) := \max_{\mathcal{U}_B} \big[H_{\mathcal{E}'}(X|E) - H_{\mathcal{E}'}(X|Y)\big], \tag{6}$$

and also find $\mathcal{U}_B$ that achieves the maximum in Eq. (6).

(ii) Find a closed-form expression of

$$F_2(\mathcal{E}) := \max_{\mathcal{U}_A, \mathcal{U}_B} \big[H_{\mathcal{E}'}(X|E) - H_{\mathcal{E}'}(X|Y)\big], \tag{7}$$

and also find $(\mathcal{U}_A, \mathcal{U}_B)$ that achieve the maximum in Eq. (7).

Eqs. (6) and (7) are the key generation rates optimized within the one-side compensation and the two-side compensation respectively.

Next, we consider the case in which Alice and Bob use the conventional channel estimation (see [26] for the detail of the conventional estimation). From Eq. (2), we find that Alice and Bob can only estimate the parameters $\gamma' = (R'_{zz}, R'_{xx}, R'_{yy})$, and they cannot estimate the parameters $\kappa' = (R'_{zx}, R'_{zy}, R'_{xz}, R'_{xy}, R'_{yz}, R'_{yx}, t'_z, t'_x, t'_y)$. Since we have to consider the worst case with respect to the parameters κ' that cannot be estimated, we consider the following two quantities:

$$\tilde{F}_1(\mathcal{E}) := \max_{\mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_s(\gamma')} \big[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\big] \tag{8}$$

and

$$\tilde{F}_2(\mathcal{E}) := \max_{\mathcal{U}_A, \mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_s(\gamma')} \big[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\big], \tag{9}$$

where $\mathcal{P}_s(\gamma')$ is the set of all channel for given γ', i.e.,

$$\mathcal{P}_s\big(\gamma'\big) := \big\{\tilde{\mathcal{E}} = (\tilde{\gamma}, \tilde{\kappa}) : \tilde{\gamma} = \gamma'\big\}.$$

Since the definition of $\tilde{F}_1(\mathcal{E})$ and $\tilde{F}_2(\mathcal{E})$ involve the minimization, we have $F_1(\mathcal{E}) \geq \tilde{F}_1(\mathcal{E})$ and $F_2(\mathcal{E}) \geq \tilde{F}_2(\mathcal{E})$ [26].

2.3. BB84 protocol

2.3.1. Compensation within z–x plane

The BB84 protocol is almost the same as the six-state protocol. However in the BB84 protocol with the axis compensation within the z–x plane, Alice uses only z basis and x basis to transmit the bit sequence, and Bob uses only observable σ_z and σ_x to receive the bit sequence. Therefore, from Eq. (2), we find that Alice and Bob can only estimate the parameters $\omega = (R_{zz}, R_{zx}, R_{xz}, R_{xx}, t_z, t_x)$, and that they cannot estimate the parameters $\tau = (R_{zy}, R_{xy}, R_{yz}, R_{yx}, R_{yy}, t_y)$. We consider the following two kinds of compensations:

(i) Only Bob is allowed to compensate his axis within the z–x plane, i.e., the channel after the compensation is given by Eq. (3), where $\mathcal{U}_B$ is a unitary channel that rotate the Bloch sphere within the z–x plane.
(ii) Both Alice and Bob are allowed to compensate their axes within the z–x plane, i.e., the channel after the compensation is given by Eq. (4), where $\mathcal{U}_A$ and $\mathcal{U}_B$ are unitary channels that rotate the Bloch sphere within the z–x plane.

Based on the estimate of the parameters ω, Bob (or both Alice and Bob) decides $\mathcal{U}_B$ (or $\mathcal{U}_A$ and $\mathcal{U}_B$), and he (they) compensates the channel. The choice of $\mathcal{U}_B$ (or $\mathcal{U}_A$ and $\mathcal{U}_B$) can be decided according to Theorem 10 and Corollary 11 of Section 3.2.1.

As in the six-state protocol, we also employ the standard postprocessing. We first consider the case in which Alice and Bob use the accurate channel estimation. Note that Alice and Bob can only estimate the parameters $\omega' = (R'_{zz}, R'_{zx}, R'_{xz}, R'_{xx}, t'_z, t'_x)$, and that they cannot estimate the parameters $\tau' = (R'_{zy}, R'_{xy}, R'_{yz}, R'_{yx}, R'_{yy}, t'_y)$.

Let $\mathcal{P}_{ba}(\omega')$ be the set of all channels for given ω', i.e.,

$$\mathcal{P}_{ba}(\omega') := \{\tilde{\mathcal{E}} = (\tilde{\omega}, \tilde{\tau}) : \tilde{\omega} = \omega'\}.$$

In the BB84 protocol, Alice and Bob can asymptotically share a secure key if the length ℓ of the key satisfies

$$\frac{\ell}{n} < \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right]$$

(see [21,26]). Therefore, we consider the following two optimization problems:

(i) Find a closed-form expression of

$$G_1(\mathcal{E}) := \max_{\mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right], \tag{10}$$

and also find $\mathcal{U}_B$ that achieves the maximum in Eq. (10), where $\mathcal{U}_B$ is a unitary channel that rotates the Bloch sphere within the z–x plane.
(ii) Find a closed-form expression of

$$G_2(\mathcal{E}) := \max_{\mathcal{U}_A, \mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right], \tag{11}$$

and also find $(\mathcal{U}_A, \mathcal{U}_B)$ that achieve the maximum in Eq. (11), where $\mathcal{U}_A$ and $\mathcal{U}_B$ are unitary channels that rotate the Bloch sphere within the z–x plane.

Next, we consider the case in which Alice and Bob use the conventional channel estimation. From Eq. (2), we find that Alice and Bob can only estimate the parameters $\mu' = (R'_{zz}, R'_{xx})$, and they cannot estimate the parameters $\nu' = (R'_{zx}, R'_{zy}, R'_{xz}, R'_{xy}, R'_{yz}, R'_{yx}, R'_{yy}, t'_z, t'_x, t'_y)$. Since we have to consider the worst case with respect to the parameters ν' that cannot be estimated, we consider the following two quantities:

$$\tilde{G}_1(\mathcal{E}) := \max_{\mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{bc}(\mu')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right], \tag{12}$$

and

$$\tilde{G}_2(\mathcal{E}) := \max_{\mathcal{U}_A, \mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{bc}(\mu')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right], \tag{13}$$

where $\mathcal{P}_{bc}(\mu')$ is the set of all channel for given μ', i.e.,

$$\mathcal{P}_{bc}(\mu') := \{\tilde{\mathcal{E}} = (\tilde{\mu}, \tilde{\nu}) : \tilde{\mu} = \mu'\}.$$

Since the range of the minimizations in the definitions of $G_1(\mathcal{E})$, $G_2(\mathcal{E})$, $\tilde{G}_1(\mathcal{E})$, and $\tilde{G}_2(\mathcal{E})$ satisfy $\mathcal{P}_{ba}(\omega') \subset \mathcal{P}_{bc}(\mu')$, we have $G_1(\mathcal{E}) \geq \tilde{G}_1(\mathcal{E})$ and $G_2(\mathcal{E}) \geq \tilde{G}_2(\mathcal{E})$ [26].

Remark 2. Although $G_1(\mathcal{E})$, $G_2(\mathcal{E})$, $\tilde{G}_1(\mathcal{E})$, and $\tilde{G}_2(\mathcal{E})$ are defined as functions of $\mathcal{E}$, it should be noted that their values on the right hand side only depend of the parameters ω, which can be estimated in the compensation procedure of this section, and do not depend on the parameters τ, which cannot be estimated. More precisely, since $\mathcal{U}_B$ is a unitary that rotates the Bloch sphere within the z–x plane, ω' only depends on ω, and we have

$$G_1\big((\omega,\tau)\big) = G_1\big((\omega,\hat{\tau})\big)$$

for any $(\omega,\hat{\tau}) \in \mathcal{P}_{ba}(\omega)$. Similar equations also hold for $G_2(\mathcal{E})$, $\tilde{G}_1(\mathcal{E})$, and $\tilde{G}_1(\mathcal{E})$. Thus, Alice and Bob can actually calculate these functions from their estimated parameters in the compensation procedure of the protocol.

2.3.2. Compensation within any direction

In this section, we consider the BB84 protocol with the axis compensation within any direction. We consider this problem because several researchers employ the compensation within any direction in the literatures [7,10,18,28,29,31,24,30,6,16].

When we employ the one-side compensation, Alice randomly sends 0 or 1 to Bob by modulating it into a transmission basis that is randomly chosen from the z-basis or the x-basis. Since Alice does not rotate the axis of the transmitter in the one-side compensation and the one-side compensation might be used to simplify the device, it seems natural that Alice only uses the z-basis and the x-basis. Bob measures received qubits by randomly using observables σ_z, σ_x or σ_y. Note that Bob can use σ_y in addition to σ_z and σ_x because he is allowed to rotate the axis of the receiver in the axis compensation phase. In this case, from Eq. (2), we find that Alice and Bob can estimate the parameters $(R_{zz}, R_{xz}, R_{yz}, R_{zx}, R_{xx}, R_{yx}, t_z, t_x, t_y)$, and they cannot estimate the parameters (R_{zy}, R_{xy}, R_{yy}). Since Bob can use σ_y, Alice and Bob can estimate (R_{yz}, R_{yx}, t_y) in addition to $(R_{zz}, R_{xz}, R_{zx}, R_{xx}, t_z, t_x)$, which can be estimated in the compensation scheme of Section 2.3.1. Based on the estimate of the parameters $(R_{zz}, R_{xz}, R_{yz}, R_{zx}, R_{xx}, R_{yx}, t_z, t_x, t_y)$, Bob decide $\mathcal{U}_B$ and compensate the channel. The choice of $\mathcal{U}_B$ can be decided according to Theorem 15 of Section 3.2.2.

On the other hand, when we employ the two-side compensation, we allow both Alice and Bob to use z-basis, x-basis, and y-basis in the axis compensation phase. In this case, from Eq. (2), we find that Alice and Bob can estimate all of the parameters (R,t). Based on the estimate of the parameters (R,t), Alice and Bob decide $\mathcal{U}_A$ and $\mathcal{U}_B$, and they compensate the channel. The choice of $\mathcal{U}_A$ and $\mathcal{U}_B$ can be decided according to Theorem 14 of Section 3.2.2.

In the bit transmission phase (after the axis compensation phase), we allow Alice and Bob to use only z-basis and x-basis. The channel estimation phase and the postprocessing phase are exactly the same as in Section 2.3.1. Note that Alice and Bob can estimate $(R'_{zz}, R'_{xz}, R'_{zx}, R'_{xx}, t'_z, t'_x)$, but they cannot estimate the other parameters, because we do not allow neither Alice nor Bob to use y-basis in the bit transmission phase. Therefore, we consider the following two optimization problems:

(i) Find a closed-form expression of

$$J_1(\mathcal{E}) := \max_{\mathcal{U}_B} \min_{\tilde{\mathcal{E}}\in\mathcal{P}_{ba}(\omega')} \big[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\big], \tag{14}$$

and also find $\mathcal{U}_B$ that achieve the maximum in Eq. (14), where $\mathcal{U}_B$ is any unitary channel.

(ii) Find a closed-form expression of

$$J_2(\mathcal{E}) := \max_{\mathcal{U}_A,\mathcal{U}_B} \min_{\tilde{\mathcal{E}}\in\mathcal{P}_{ba}(\omega')} \big[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\big], \tag{15}$$

and also find $(\mathcal{U}_A,\mathcal{U}_B)$ that achieve the maximum in Eq. (15), where $\mathcal{U}_A$ and $\mathcal{U}_B$ are any unitary channels.

We also treat the case in which Alice and Bob use the conventional channel estimation. In this case, we consider the following two quantities:

$$\tilde{J}_1(\mathcal{E}) := \max_{\mathcal{U}_B} \min_{\tilde{\mathcal{E}}\in\mathcal{P}_{bc}(\mu')} \big[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\big], \tag{16}$$

and

$$\tilde{J}_2(\mathcal{E}) := \max_{\mathcal{U}_A,\mathcal{U}_B} \min_{\tilde{\mathcal{E}}\in\mathcal{P}_{bc}(\mu')} \big[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\big], \tag{17}$$

where $\mathcal{U}_A$ and $\mathcal{U}_B$ are any unitary channels. Since the range of the minimizations in the definitions of $J_1(\mathcal{E})$, $J_2(\mathcal{E})$, $\tilde{J}_1(\mathcal{E})$, and $\tilde{J}_2(\mathcal{E})$ satisfy $\mathcal{P}_{ba}(\omega') \subset \mathcal{P}_{bc}(\mu')$, we have $J_1(\mathcal{E}) \geq \tilde{J}_1(\mathcal{E})$ and $J_2(\mathcal{E}) \geq \tilde{J}_2(\mathcal{E})$ [26].

Remark 3. Although $J_1(\mathcal{E})$ and $\tilde{J}_1(\mathcal{E})$ are defined as functions of $\mathcal{E}$, it should be noted that their values on the right hand side only depend of the parameters $(R_{zz}, R_{xz}, R_{yz}, R_{zx}, R_{xx}, R_{yx}, t_z, t_x, t_y)$, which can be estimated in the one-side compensation procedure of this section, and do not depend on the parameters (R_{zy}, R_{xy}, R_{yy}), which cannot be estimated. Thus, Alice and Bob can actually calculate these functions from their estimated parameters in the compensation procedure of the protocol.

3. Optimal compensation for unital channels

In this section, we solve the problems formulated in Sections 2.2, 2.3.1, and 2.3.2 respectively for unital channels.

3.1. Six-state protocol

For any channel $\mathcal{E} = (R, t)$, by the singular value decomposition, we can decompose[2] the matrix R as

$$\begin{aligned} R &= B\,\mathrm{diag}[e_z, e_x, e_y]\,A \\ &= \begin{bmatrix} \langle B_z| \\ \langle B_x| \\ \langle B_y| \end{bmatrix} \begin{bmatrix} e_z & 0 & 0 \\ 0 & e_x & 0 \\ 0 & 0 & e_y \end{bmatrix} \begin{bmatrix} |A_z\rangle & |A_x\rangle & |A_y\rangle \end{bmatrix} \\ &= \begin{bmatrix} \langle B_z|\tilde{A}_z\rangle & \langle B_z|\tilde{A}_x\rangle & \langle B_z|\tilde{A}_y\rangle \\ \langle B_x|\tilde{A}_z\rangle & \langle B_x|\tilde{A}_x\rangle & \langle B_x|\tilde{A}_y\rangle \\ \langle B_y|\tilde{A}_z\rangle & \langle B_y|\tilde{A}_x\rangle & \langle B_y|\tilde{A}_y\rangle \end{bmatrix}, \end{aligned} \tag{18}$$

where A and B are the rotation matrices[3], $|A_z\rangle$, $|A_x\rangle$, and $|A_y\rangle$ are the column vectors of A with norm 1, $\langle B_z|$, $\langle B_x|$, and $\langle B_y|$ are the row vectors of B with norm 1, $|e_z|$, $|e_x|$, and $|e_y|$ are the singular values of R, and we set $\langle \tilde{A}_z| = (e_z A_{zz}, e_x A_{zx}, e_y A_{zy})$, $\langle \tilde{A}_x| = (e_z A_{xz}, e_x A_{xx}, e_y A_{xy})$, and $\langle \tilde{A}_y| = (e_z A_{yz}, e_x A_{yx}, e_y A_{yy})$.

Henceforth, we identify Alice's compensation $\mathcal{U}_A$ and Bob's compensation $\mathcal{U}_B$ with the 3×3 rotation matrices O_A and O_B. Then, the matrix part of the Stokes parameterization of the compensated channel $\mathcal{E}' = (R', t')$ is given by $R' = O_B R O_A$.

The following theorem gives a closed-form expression of the key generation rate optimized by the two-side compensation.

Theorem 4. *Suppose that $\mathcal{E}$ is a unital channel. Let $O_A^* = A^{-1}$ and $O_B^* = B^{-1}$, and let $\mathcal{U}_A^*$ and $\mathcal{U}_B^*$ be the unitary channels corresponding to O_A^* and O_B^* respectively. Then, the compensated channel $\mathcal{E}^* = \mathcal{U}_B^* \circ \mathcal{E} \circ \mathcal{U}_A^*$ is the Pauli channel such that the matrix part of the Stokes parameterization is given by $R^* = diag[e_z, e_x, e_y]$, and $\mathcal{E}^*$ satisfies*

$$F_2(\mathcal{E}) = \max_{\mathcal{U}_A, \mathcal{U}_B} \left[H_{\mathcal{E}'}(X|E) - H_{\mathcal{E}'}(X|Y) \right] \tag{19}$$

$$= H_{\mathcal{E}^*}(X|E) - H_{\mathcal{E}^*}(X|Y) \tag{20}$$

$$= 1 - H[q_i, q_z, q_x, q_y], \tag{21}$$

where $H[q_i, q_z, q_x, q_y]$ is the Shannon entropy [9] of the distribution

$$q_i = \frac{1 + e_z + e_x + e_y}{4}, \tag{22}$$

$$q_z = \frac{1 + e_z - e_x - e_y}{4}, \tag{23}$$

$$q_x = \frac{1 - e_z + e_x - e_y}{4}, \tag{24}$$

$$q_y = \frac{1 - e_z - e_x + e_y}{4}. \tag{25}$$

Furthermore, the maximum in Eq. (19) is achieved without any compensation, i.e.,

$$H_{\mathcal{E}}(X|E) - H_{\mathcal{E}}(X|Y) = H_{\mathcal{E}^*}(X|E) - H_{\mathcal{E}^*}(X|Y)$$

if and only if the vectors $|\tilde{A}_z\rangle$ and $|B_z\rangle$ are scalar multiple of each other. ■

The first statement implies that an optimal compensation procedure is to compensate the channel to a Pauli channel. The second statement implies that $(\mathcal{U}_A, \mathcal{U}_B)$ achieving the maximum are not unique.

[2] The decomposition is not unique because we can change the order of (e_z, e_x, e_y) or the sign of them by adjusting the rotation matrices A and B. However, the result in this paper does not depends on a choice of the decomposition.

[3] The rotation matrix is the real orthogonal matrix with determinant 1.

Remark 5. Although the channel $\mathcal{E}$ is assumed to be a unital channel in Theorem 4, it should be noted that our protocol is secure for any channel. Theorem 4 is just saying that the compensation (O_A^*, O_B^*) given in this theorem is optimal if the channel is a unital channel. Similar remarks are also applied for the other theorems and corollaries throughout the paper. ■

The following corollary gives a closed-form expression of the key generation rate optimized by the one-side compensation.

Corollary 6. *Suppose that $\mathcal{E}$ is a unital channel. Let*

$$O_B^* = \begin{bmatrix} \langle O_{B,\mathsf{z}}^*| \\ \langle O_{B,\mathsf{x}}^*| \\ \langle O_{B,\mathsf{y}}^*| \end{bmatrix}$$

be a rotation matrix such that $\langle O_{B,\mathsf{z}}^*|$ is a scalar multiple of $(R_{\mathsf{zz}}, R_{\mathsf{xz}}, R_{\mathsf{yz}})$, where $\langle O_{B,\mathsf{x}^*}|$ and $\langle O_{B,\mathsf{y}}^*|$ can be arbitrary as long as they constitute a rotation matrix, and let $\mathcal{U}_B^*$ the unitary channel corresponding to O_B^*. Then, the compensated channel $\mathcal{E}^* = \mathcal{U}_B^* \circ \mathcal{E}$ satisfies

$$\begin{aligned} F_1(\mathcal{E}) &= H_{\mathcal{E}^*}(X|E) - H_{\mathcal{E}^*}(X|Y) \\ &= F_2(\mathcal{E}). \end{aligned}$$ ■

Note that Corollary 6 follows from the second statement of Theorem 4.

Surprisingly, we do not lose any optimality even if we only allow Bob to compensate his axis (one-side compensation). This fact is useful to simplify the implementation of the optimal compensation procedure.

Since $H_{\mathcal{E}}(X|Y) = h((1+R_{\mathsf{zz}})/2)$ for any unital channel and $R_{\mathsf{zz}} = \langle B_{\mathsf{z}}|\tilde{A}_{\mathsf{z}}\rangle$, we find that an optimal one-side compensation procedure is to compensate the channel so that Bob can detect Alice's transmitted state most reliably, i.e., $H_{\mathcal{E}'}(X|Y)$ is minimized, where $h(\cdot)$ is the binary entropy function. Note that the fact that $|\tilde{A}_{\mathsf{z}}\rangle$ and $|B'_{\mathsf{z}}\rangle$ are scalar multiple of each other does not necessarily mean the compensated channel $\mathcal{E}'$ is a Pauli channel.

Next, we consider the case in which Alice and Bob use the conventional channel estimation. The following theorem states that the optimized key generation rate with the accurate channel estimation coincides with that with the conventional channel estimation if we use the two-side compensation. The following theorem also gives the necessary and sufficient condition such that the optimized key generation rates with the accurate channel estimation and the conventional channel estimation coincide when we use the one-side compensation.

Theorem 7. *Suppose that $\mathcal{E}$ is a unital channel. Then, we have*

$$F_2(\mathcal{E}) = \tilde{F}_2(\mathcal{E}),$$

where $\tilde{F}_2(\mathcal{E})$ is achieved by O_A^* and O_B^* specified in Theorem 4. Furthermore, we have

$$F_1(\mathcal{E}) = \tilde{F}_1(\mathcal{E})$$

if and only if $|\tilde{A}_{\mathsf{z}}\rangle$, $|\tilde{A}_{\mathsf{x}}\rangle$, and $|\tilde{A}_{\mathsf{y}}\rangle$ are orthogonal to each other. If this condition is satisfied, then $\tilde{F}_1(\mathcal{E})$ is achieved by O_B^* such that $\langle O_{B,\mathsf{z}}^*|$ and $\langle O_{B,\mathsf{x}}^*|$ and $\langle O_{B,\mathsf{y}}^*|$ are scalar multiple of $(R_{\mathsf{zz}}, R_{\mathsf{xz}}, R_{\mathsf{yz}})$, $(R_{\mathsf{zx}}, R_{\mathsf{xx}}, R_{\mathsf{yx}})$, and $(R_{\mathsf{zy}}, R_{\mathsf{xy}}, R_{\mathsf{yy}})$ respectively. ■

Corollary 8. *Suppose that $\mathcal{E}$ is a unital channel. Then, we have*

$$\tilde{F}_1(\mathcal{E}) = \tilde{F}_2(\mathcal{E})$$

if and only if $|\tilde{A}_{\mathsf{z}}\rangle$, $|\tilde{A}_{\mathsf{x}}\rangle$, and $|\tilde{A}_{\mathsf{y}}\rangle$ are orthogonal to each other. ■

3.2. BB84 protocol

3.2.1. Compensation within z–x plane

For any channel $\mathcal{E} = (R, t)$, by the singular value decomposition, we can decompose the left upper 2×2 sub-matrix S of the matrix R as

$$\begin{aligned} S &= V \operatorname{diag}[d_{\mathsf{z}}, d_{\mathsf{x}}]\, U \\ &= \begin{bmatrix} \langle V_{\mathsf{z}}| \\ \langle V_{\mathsf{x}}| \end{bmatrix} \begin{bmatrix} d_{\mathsf{z}} & 0 \\ 0 & d_{\mathsf{x}} \end{bmatrix} \begin{bmatrix} |U_{\mathsf{z}}\rangle & |U_{\mathsf{x}}\rangle \end{bmatrix} \\ &= \begin{bmatrix} \langle V_{\mathsf{z}}|\tilde{U}_{\mathsf{z}}\rangle & \langle V_{\mathsf{z}}|\tilde{U}_{\mathsf{x}}\rangle \\ \langle V_{\mathsf{x}}|\tilde{U}_{\mathsf{z}}\rangle & \langle V_{\mathsf{x}}|\tilde{U}_{\mathsf{x}}\rangle \end{bmatrix}, \end{aligned}$$

where U and V are the rotation matrices, $|d_z|$ and $|d_x|$ are the singular values of S, and we set $\langle \tilde{U}_z| = (d_z U_{zz}, d_x U_{zx})$ and $\langle \tilde{U}_x| = (d_z U_{xz}, d_x U_{xx})$.

Henceforth, we identify Alice's compensation $\mathcal{U}_A$ and Bob's compensation $\mathcal{U}_B$ with the 2×2 rotation matrices Q_A and Q_B, because their compensation are restricted within the z–x plane. Note that the left upper 2×2 sub-matrix S' of the matrix R' of the compensated channel is given by $S' = Q_B S Q_A$.

The following lemma provides a closed-form expression of the key generation rate with the accurate channel estimation for unital channels, and it will be used several times in the rest of this paper.

Lemma 9. *For any unital channel $\mathcal{E} = (\omega, \tau)$, we have*

$$\begin{aligned}&\min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \\ &= 1 - h\left(\frac{1+d_z}{2}\right) - h\left(\frac{1+d_x}{2}\right) + h\left(\frac{1+\sqrt{R_{zz}^2 + R_{xz}^2}}{2}\right) - h\left(\frac{1+R_{zz}}{2}\right).\end{aligned}$$

Proof of Lemma 9. This lemma follows from [26, Proposition 2] and the fact $H_{\mathcal{E}}(X|Y) = h((1 + R_{zz})/2)$ for any unital channel. □

The following theorem gives a closed-form expression of the key generation rate optimized by the two-side compensation.

Theorem 10. *Suppose that $\mathcal{E}$ is a unital channel. Let $Q_A^* = U^{-1}$ and $Q_B^* = V^{-1}$, and let $\mathcal{U}_A^*$ and $\mathcal{U}_B^*$ be the unitary channels corresponding to Q_A^* and Q_B^* respectively. Then, the compensated channel $\mathcal{U}_B^* \circ \mathcal{E} \circ \mathcal{U}_A^* =: \mathcal{E}^* = (\omega^*, \tau^*)$ satisfies*

$$G_2(\mathcal{E}) = \max_{\mathcal{U}_A, \mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \tag{26}$$

$$= \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega^*)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \tag{27}$$

$$= 1 - h\left(\frac{1+d_z}{2}\right) - h\left(\frac{1+d_x}{2}\right). \tag{28}$$

Furthermore, the maximum is achieved without any compensation, i.e.,

$$\min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] = \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega^*)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right]$$

if and only if the vectors $|\tilde{U}_z\rangle$ and $|V_z\rangle$ are scalar multiple of each other. ■

The first statement implies that an optimal compensation procedure is to compensate the channel to a channel such that the left upper sub-matrix S' of the Stokes parameterization of the compensated channel is a diagonal matrix. The latter statement implies that $(\mathcal{U}_A, \mathcal{U}_B)$ achieving the maximum is not unique.

By using Theorem 10, we can derive the following corollary, which gives the key generation rate optimized by the one-side compensation.

Corollary 11. *Suppose that $\mathcal{E}$ is a unital channel. Let*

$$Q_B^* = \begin{bmatrix} \langle Q_{B,z}^*| \\ \langle Q_{B,x}^*| \end{bmatrix}$$

be a rotation matrix such that $\langle Q_{B,z}^*|$ is a scalar multiple of (R_{zz}, R_{xz}),[4] and let $\mathcal{U}_B^*$ be the unitary channel corresponding to O_B^*. Then, the compensated channel $\mathcal{U}_B^* \circ \mathcal{E} =: \mathcal{E}^* = (\omega^*, \tau^*)$ satisfies

$$G_1(\mathcal{E}) = \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega^*)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \tag{29}$$

$$= G_2(\mathcal{E}). \quad ■ \tag{30}$$

[4] Note that $\langle Q_{B,x}|$ is uniquely determined from $\langle Q_{B,z}|$ because they constitute a rotation matrix.

Note that Corollary 11 follows from the latter statement of Theorem 10.

Surprisingly, we do not lose any optimality even if we only allow Bob to compensate his axis (one-side compensation). This fact is useful to simplify the implementation of the optimal compensation procedure.

Since $H_{\mathcal{E}}(X|Y) = h((1+R_{zz})/2)$ for any unital channel and $R_{zz} = \langle V_z|\tilde{U}_z\rangle$, we find that an optimal one-side compensation procedure is to compensate the channel so that Bob can detect Alice's transmitted state most reliably, i.e., $H_{\mathcal{E}'}(X|Y)$ is minimized. Note that the fact that $|\tilde{U}_z\rangle$ and $|V'_z\rangle$ are scalar multiple of each other does not necessarily mean that the left upper sub-matrix S' of the Stokes parameterization of the compensated channel is a diagonal matrix.

Next, we consider the case in which Alice and Bob use the conventional channel estimation. The following theorem states that the optimized key generation rate with the accurate channel estimation coincides with that with the conventional channel estimation if we use the two-side compensation. The following theorem also gives the necessary and sufficient condition such that the optimized key generation rates with the accurate channel estimation and the conventional channel estimation coincide when we use the one-side compensation.

Theorem 12. *Suppose that $\mathcal{E}$ is a unital channel. Then, we have*

$$G_2(\mathcal{E}) = \tilde{G}_2(\mathcal{E}),$$

where $\tilde{G}_2(\mathcal{E})$ is achieved by Q_A^* and Q_B^* specified in Theorem 10. Furthermore, we have

$$G_1(\mathcal{E}) = \tilde{G}_1(\mathcal{E})$$

if and only if $|\tilde{U}_z\rangle$ and $|\tilde{U}_x\rangle$ are orthogonal to each other. If this condition is satisfied, $\tilde{G}_1(\mathcal{E})$ is achieved by Q_B^* such that $\langle Q_{B,z}^*|$ and $\langle Q_{B,x}^*|$ are scalar multiple of (R_{zz}, R_{xz}) and (R_{zx}, R_{xx}) respectively. ■

Corollary 13. *Suppose that $\mathcal{E}$ is a unital channel. Then, we have*

$$\tilde{G}_1(\mathcal{E}) = \tilde{G}_2(\mathcal{E})$$

if and only if $|\tilde{U}_z\rangle$ and $|\tilde{U}_x\rangle$ are orthogonal to each other. ■

3.2.2. Compensation within any direction

In this section, we consider the case in which either Alice or Bob are allowed to compensate their axes within any direction [7,10,18,28,29,31,24,30,6,16]. For any channel $\mathcal{E} = (R, t)$, by the singular value decomposition, we can decompose the matrix R as in Eq. (18). Furthermore, we identify Alice's compensation $\mathcal{U}_A$ and Bob's compensation $\mathcal{U}_B$ with the 3×3 rotation matrices O_A and O_B as in Section 3.1. When we consider the compensation within any direction, it should be noted that we can estimate all the parameters in the two-side compensation and only a part of the parameters in the one-side compensation (see also Section 2.3.2).

The following theorem gives the key generation rate optimized by the two-side compensation.

Theorem 14. *Suppose that $\mathcal{E}$ is a unital channel. Let $\mathcal{U}_A^*$ and $\mathcal{U}_B^*$ be unitary channels such that the compensated channel $\mathcal{U}_B^* \circ \mathcal{E} \circ \mathcal{U}_A^* =: \mathcal{E}^* = (\omega^*, \tau^*)$ is a Pauli channel and the singular values $|e_z^*|$, $|e_x^*|$, and $|e_y^*|$ of $R^* = \mathrm{diag}[e_z^*, e_x^*, e_y^*]$ satisfy*

$$|e_z^*| \geq |e_x^*| \geq |e_y^*|.$$

Then, we have

$$J_2(\mathcal{E}) = \max_{\mathcal{U}_A, \mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \tag{31}$$

$$= \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega^*)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \tag{32}$$

$$= 1 - h\left(\frac{1+e_z^*}{2}\right)s - h\left(\frac{1+e_x^*}{2}\right). \quad \blacksquare \tag{33}$$

The following theorem gives the key generation rate optimized by the one-side compensation.

Theorem 15. *Suppose that $\mathcal{E}$ is a unital channel. Let*

$$O_B^* = \begin{bmatrix} \langle O_{B,z}^*| \\ \langle O_{B,x}^*| \\ \langle O_{B,y}^*| \end{bmatrix}$$

be a rotation matrix such that $\langle O^*_{B,z}|$ *and* $\langle O^*_{B,x}|$ *span the same subspace as that spanned by* (R_{zz}, R_{xz}, R_{yz}) *and* (R_{zx}, R_{xx}, R_{yx})*, and that* $\langle O^*_{B,z}|$ *is a scalar multiple of* (R_{zz}, R_{xz}, R_{yz})*,*[5] *and let* $\mathcal{U}^*_B$ *be the unitary channel corresponding to* O^*_B*. Then, the compensated channel* $\mathcal{U}^*_B \circ \mathcal{E} =: \mathcal{E}^* = (\omega^*, \tau^*)$ *satisfies*

$$J_1(\mathcal{E}) = \max_{\mathcal{U}_B} \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega')} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y) \right] \tag{34}$$

$$= \min_{\tilde{\mathcal{E}} \in \mathcal{P}_{ba}(\omega^*)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y) \right] \tag{35}$$

$$= 1 - h\left(\frac{1+s^*_1}{2} \right) - h\left(\frac{1+s^*_2}{2} \right), \tag{36}$$

where s^*_1 *and* s^*_2 *are the singular values of the upper left* 2×2 *sub-matrix matrix*

$$S^* = \begin{bmatrix} \langle B^*_z|\tilde{A}_z\rangle & \langle B^*_z|\tilde{A}_x\rangle \\ \langle B^*_x|\tilde{A}_z\rangle & \langle B^*_x|\tilde{A}_x\rangle \end{bmatrix}$$

of $R^* = O^*_B R$ *such that* $s^*_1 \geq s^*_2$*.* ■

Remark 16. The equality

$$J_1(\mathcal{E}) = J_2(\mathcal{E})$$

does not hold in general. For example, $J_1(\mathcal{E}) \neq J_2(\mathcal{E})$ if $R = \mathrm{diag}[e_z, e_x, e_y]$ and $|e_z| < |e_x| < |e_y|$. ■

Next, we consider the case in which Alice and Bob use the conventional channel estimation. The following theorem states that the optimized key generation rate with the accurate channel estimation coincides with that with the conventional channel estimation if we use the two-side compensation. The following theorem also gives the necessary and sufficient condition such that the optimized key generation rates with the accurate channel estimation and the conventional channel estimation coincide.

Theorem 17. *Suppose that* $\mathcal{E}$ *is a unital channel. Then, we have*

$$J_2(\mathcal{E}) = \tilde{J}_2(\mathcal{E}),$$

where $\tilde{J}_2(\mathcal{E})$ *is achieved by* O^*_A *and* O^*_B *specified in Theorem 14. Furthermore, we have*

$$J_1(\mathcal{E}) = \tilde{J}_1(\mathcal{E})$$

if and only if $|\tilde{A}_z\rangle$ *and* $|\tilde{A}_x\rangle$ *are orthogonal to each other. If this condition is satisfied, then* $\tilde{J}_1(\mathcal{E})$ *is achieved by* O^*_B *such that* $\langle O^*_{B,z}|$ *and* $\langle O^*_{B,x}|$ *are scalar multiple of* (R_{zz}, R_{xz}, R_{yz}) *and* (R_{zx}, R_{xx}, R_{yx}) *respectively.* ■

Proof of Theorem 17. This theorem can be proved almost in a similar manner to Theorem 12. Therefore, we omit the proof. □

Corollary 18. *Suppose that* $\mathcal{E}$ *is a unital channel. Then, we have*

$$\tilde{J}_1(\mathcal{E}) < \tilde{J}_2(\mathcal{E})$$

if $|\tilde{A}_z\rangle$ *and* $|\tilde{A}_x\rangle$ *are not orthogonal to each other.* ■

4. Conclusion

In this paper, we investigated the axis compensation in the QKD protocols in various settings. We clarified optimal compensation procedures over unital channels for one-side compensation with the accurate channel estimation and for two-side compensation with both estimations, i.e., the conventional channel estimation and the accurate channel estimation, while we could not identify an optimal compensation procedure for one-side compensation with the conventional channel estimation. Although our proposed compensation procedures are optimal for unital channels, it is not clear whether those compensation procedures are optimal or not for general channels. We also clarified that the optimized key generation rates with the conventional channel estimation are strictly smaller than the optimized key generation rates with the accurate

[5] Note that $\langle O_{B,y}|$ is uniquely determined from $\langle O_{B,z}|$ and $\langle O_{B,x}|$ because they constitute a rotation matrix.

channel estimation for the one-side compensation. Our results imply that we should use the accurate channel estimation when we employ the one-side compensation. On the other hand, we do not have to use the accurate channel estimation when we employ the two-side compensation.

Although we clarified the optimal compensation procedures for the standard postprocessing, it is an important future research agenda to clarify the optimal compensation procedures when we employ more complicated postprocessing (e.g. the postprocessing with the noisy preprocessing [22,15] or the two-way classical communication [14,27]).

Acknowledgements

The authors would like to thank Dr. Toyohiro Tsurumaru for bringing the axis compensation problem to our attention. The first author would like to thank Prof. Yasutada Oohama for his support. The authors also would like to thank the editors of this special issue, Prof. Renato Renner and Prof. Tal Mor, and reviewers for valuable comments. This research is partly supported by the Japan Society of Promotion of Science under Grants-in-Aid No. 00197137.

Appendix A. Proof of Theorem 4

The equality between Eqs. (20) and (21) is well known (e.g. see [22] or [26, Eq. (20)]). Since Eq. (19) is obviously larger than or equals to Eq. (20), it suffices to show that Eq. (19) is smaller than or equals to Eq. (21) for any $\mathcal{U}_A$ and $\mathcal{U}_B$. For any fixed $\mathcal{U}_A$ and $\mathcal{U}_B$, by using [26, Eq. (20)] and the discussions right before it, Eq. (19) can be rewritten as

$$1 - H[q_{\mathsf{i}}, q_{\mathsf{z}}, q_{\mathsf{x}}, q_{\mathsf{y}}] + h\left(\frac{1 + \||\tilde{A}'_{\mathsf{z}}\rangle\|}{2}\right) - h\left(\frac{1 + \langle B'_{\mathsf{z}}|\tilde{A}'_{\mathsf{z}}\rangle}{2}\right).$$

From the form of $h(\cdot)$, Cauchy's inequality $|\langle B'_{\mathsf{z}}|\tilde{A}'_{\mathsf{z}}\rangle| \le \||\tilde{A}'_{\mathsf{z}}\rangle\|$ implies that Eq. (19) is smaller than or equals to Eq. (21). The equality holds if and only if the vectors $|\tilde{A}'_{\mathsf{z}}\rangle$ and $|B'_{\mathsf{z}}\rangle$ are scalar multiple of each other, which is exactly the second statement of the theorem. □

Appendix B. Proof of Theorem 7

Let $\mathcal{E}^*$ be the Pauli channel defined in Theorem 4. Then, we have

$$\begin{aligned} F_2(\mathcal{E}) &\ge \tilde{F}_2(\mathcal{E}) \\ &\ge \min_{\tilde{\mathcal{E}} \in \mathcal{P}_s(\gamma^*)} \left[H_{\tilde{\mathcal{E}}}(X|E) - H_{\tilde{\mathcal{E}}}(X|Y)\right] \\ &= 1 - H[q_{\mathsf{i}}, q_{\mathsf{z}}, q_{\mathsf{x}}, q_{\mathsf{y}}] \\ &= F_2(\mathcal{E}), \end{aligned}$$

which implies the first statement of the theorem.

To prove the "if" part of the second statement, assume that $|\tilde{A}_{\mathsf{z}}\rangle$, $|\tilde{A}_{\mathsf{x}}\rangle$, and $|\tilde{A}_{\mathsf{y}}\rangle$ are orthogonal to each other. Then, we can take a rotation matrix O_B^* so that $\langle O^*_{B,\mathsf{z}}|$ and $\langle O^*_{B,\mathsf{x}}|$ and $\langle O^*_{B,\mathsf{y}}|$ are scalar multiple of $(R_{\mathsf{zz}}, R_{\mathsf{xz}}, R_{\mathsf{yz}})$, $(R_{\mathsf{zx}}, R_{\mathsf{xx}}, R_{\mathsf{yx}})$, and $(R_{\mathsf{zy}}, R_{\mathsf{xy}}, R_{\mathsf{yy}})$ respectively, and we have $R' = O_B^* R = \mathrm{diag}[e_{\mathsf{z}}, e_{\mathsf{x}}, e_{\mathsf{y}}]$. Thus, we have

$$F_1(\mathcal{E}) \ge \tilde{F}_1(\mathcal{E}) \ge 1 - H[q_{\mathsf{i}}, q_{\mathsf{z}}, q_{\mathsf{x}}, q_{\mathsf{y}}] = F_1(\mathcal{E}).$$

Next, we show the "only if" part of the second statement. Suppose that at least one pair of $|\tilde{A}_{\mathsf{z}}\rangle$, $|\tilde{A}_{\mathsf{x}}\rangle$, and $|\tilde{A}_{\mathsf{y}}\rangle$ is not orthogonal to each other. Then, for arbitrarily fixed $\mathcal{U}_B$, the compensated channel $\mathcal{E}'$ is not a Pauli channel, i.e., the Choi operator ρ' is not a Bell diagonal state. Let $\rho'_{\mathsf{a}} := (\bar{\sigma}_{\mathsf{a}} \otimes \sigma_{\mathsf{a}})\rho'(\bar{\sigma}_{\mathsf{a}} \otimes \sigma_{\mathsf{a}})$ for $\mathsf{a} \in \{\mathsf{i}, \mathsf{z}, \mathsf{x}, \mathsf{y}\}$, where $\bar{\sigma}_{\mathsf{a}}$ is the complex conjugate of σ_{a}. Since ρ' is not Bell diagonal state, at least one of ρ'_{z}, ρ'_{x}, and ρ'_{y} is different from ρ'_{i}. Let

$$\rho^{tw} := \sum_{\mathsf{a} \in \{\mathsf{i},\mathsf{z},\mathsf{x},\mathsf{y}\}} \frac{1}{4}\rho'_{\mathsf{a}}$$

be the partially twirled state [2]. Then, since the von Neumann entropy is a strict concave function [19], we have

$$\begin{aligned} \tilde{F}_1(\mathcal{E}) &= \max_{\mathcal{U}_B}\left[1 - H\left(\rho^{tw}\right)\right] \\ &< \max_{\mathcal{U}_B}\left[1 - \sum_{\mathsf{a} \in \{\mathsf{i},\mathsf{z},\mathsf{x},\mathsf{y}\}} \frac{1}{4} H\left(\rho'_{\mathsf{a}}\right)\right] \\ &= \max_{\mathcal{U}_B}\left[1 - H\left(\rho'\right)\right] \\ &= F_1(\mathcal{E}). \qquad \square \end{aligned}$$

Appendix C. Proof of Theorem 10

By using Lemma 9, we have the equality between Eqs. (27) and (28). Since Eq. (26) is obviously larger than or equals to Eq. (27), it suffices to show that Eq. (26) is smaller than or equals to Eq. (28). For any fixed $\mathcal{U}_A$ and $\mathcal{U}_B$, by using Lemma 9 again, Eq. (26) can be rewritten as

$$1-h\left(\frac{1+d_z}{2}\right)-h\left(\frac{1+d_x}{2}\right)+h\left(\frac{1+\||\tilde{U}'_z\rangle\|}{2}\right)-h\left(\frac{1+\langle V'_z|\tilde{U}'_z\rangle}{2}\right).$$

From the form of $h(\cdot)$, Cauchy's inequality $|\langle V'_z|\tilde{U}'_z\rangle| \le \||\tilde{U}'_z\rangle\|$ implies that Eq. (26) is smaller than or equals to Eq. (28). The equality holds if and only if the vectors $|\tilde{U}'_z\rangle$ and $|V'_z\rangle$ are scalar multiple of each other, which is exactly the second statement of the theorem. □

Appendix D. Proof of Theorem 12

Let $\mathcal{E}^*$ be the channel defined in Theorem 10. Then, we have

$$\begin{aligned}
G_2(\mathcal{E}) &\ge \tilde{G}_2(\mathcal{E}) \\
&\ge \min_{\tilde{\mathcal{E}}\in\mathcal{P}_{bc}(\mu^*)} \left[H_{\tilde{\mathcal{E}}}(X|E)-H_{\tilde{\mathcal{E}}}(X|Y)\right] \\
&= 1-h\left(\frac{1+d_z}{2}\right)-h\left(\frac{1+d_x}{2}\right) \\
&= G_2(\mathcal{E}),
\end{aligned}$$

which implies the first statement of the theorem.

To prove the "if" part of the second statement, assume that $|\tilde{U}_z\rangle$ and $|\tilde{U}_x\rangle$ are orthogonal to each other. Then, we can take a rotation matrix Q_B^* so that $\langle Q^*_{B,z}|$ and $\langle Q^*_{B,x}|$ are scalar multiple of (R_{zz}, R_{xz}) and (R_{zx}, R_{xx}) respectively, and we have $S' = Q_B^* S = \mathrm{diag}[d_z, d_x]$. Then, we have

$$G_1(\mathcal{E}) \ge \tilde{G}_1(\mathcal{E}) \ge 1-h\left(\frac{1+d_z}{2}\right)-h\left(\frac{1+d_x}{2}\right) = G_1(\mathcal{E}).$$

Next, we show the "only if" part. Suppose that $|\tilde{U}_z\rangle$ and $|\tilde{U}_x\rangle$ are not orthogonal to each other. Then, for an arbitrarily fixed $\mathcal{U}_B$, either $\langle V'_z|\tilde{U}_x\rangle \neq 0$ or $\langle V'_x|\tilde{U}_z\rangle \neq 0$ holds. Then, we have

$$\begin{aligned}
\tilde{G}_1(\mathcal{E}) &= \max_{\mathcal{U}_B}\left[1-h\left(\frac{1+\langle V'_z|\tilde{U}_z\rangle}{2}\right)-h\left(\frac{1+\langle V'_x|\tilde{U}_x\rangle}{2}\right)\right] \\
&< 1-h\left(\frac{1+\||\tilde{U}_z\rangle\|}{2}\right)-h\left(\frac{1+\||\tilde{U}_x\rangle\|}{2}\right) \\
&= 1-h\left(\frac{1+\sqrt{d_z^2U_{zz}^2+d_x^2U_{zx}^2}}{2}\right)-h\left(\frac{1+\sqrt{d_z^2U_{xz}^2+d_x^2U_{xx}^2}}{2}\right) \\
&\le 1-U_{zz}^2h\left(\frac{1+\sqrt{d_z^2}}{2}\right)-U_{zx}^2h\left(\frac{1+\sqrt{d_x^2}}{2}\right) \\
&\quad -U_{xz}^2h\left(\frac{1+\sqrt{d_z^2}}{2}\right)-U_{xx}^2h\left(\frac{1+\sqrt{d_x^2}}{2}\right) \\
&= 1-(U_{zz}^2+U_{xz}^2)h\left(\frac{1+d_z}{2}\right)-(U_{zx}^2+U_{xx}^2)h\left(\frac{1+d_x}{2}\right) \\
&= 1-h\left(\frac{1+d_z}{2}\right)-h\left(\frac{1+d_x}{2}\right) \\
&= G_1(\mathcal{E}),
\end{aligned} \tag{D.1}$$

where we used the concavity of the function

$$h\left(\frac{1+\sqrt{x}}{2}\right) \tag{D.2}$$

in the inequality of Eq. (D.1). We can show the concavity of Eq. (D.2) by showing that the second derivative is non-positive. □

Appendix E. Proof of Theorem 14

By using Lemma 9, we have the equality between Eqs. (32) and (33). Since Eq. (31) is obviously larger than or equals to Eq. (32), it suffices to show that Eq. (31) is smaller than or equals to Eq. (33).

For any fixed $\mathcal{U}_A$ and $\mathcal{U}_B$, Theorem 10 implies

$$\begin{aligned}&\min_{\tilde{\mathcal{E}}\in\mathcal{P}_{ba}(\omega')}\left[H_{\tilde{\mathcal{E}}}(X|E)-H_{\tilde{\mathcal{E}}}(X|Y)\right]\\&\le G_2\left(\mathcal{E}'\right)\\&=1-h\left(\frac{1+d'_{\mathsf{z}}}{2}\right)-h\left(\frac{1+d'_{\mathsf{x}}}{2}\right),\end{aligned}\tag{E.1}$$

where $|d'_{\mathsf{z}}|$ and $|d'_{\mathsf{x}}|$ are the singular values of the left upper 2×2 sub-matrices S' of R' of the compensated channel $\mathcal{E}'$.

Note that the singular values of R' are equal to those of R^*. By using the interlacing inequalities for singular values of sub-matrices [23], we have

$$|e^*_{\mathsf{z}}|\ge\max\left[|d'_{\mathsf{z}}|,|d'_{\mathsf{x}}|\right]$$

and

$$|e^*_{\mathsf{x}}|\ge\min\left[|d'_{\mathsf{z}}|,|d'_{\mathsf{x}}|\right].$$

These inequalities imply that Eq. (E.1) is smaller than or equals to Eq. (33), which completes the proof. □

Appendix F. Proof of Theorem 15

The second statement of Theorem 10 implies that the equality between Eqs. (35) and (36). Since Eq. (34) is obviously larger than or equals to Eq. (35), we show that Eq. (34) is smaller than or equals to Eq. (36).

For arbitrarily fixed O_B, let s'_1 and s'_2 be the singular values of the upper left 2×2 matrix

$$S'=\begin{bmatrix}\langle B'_{\mathsf{z}}|\tilde{A}_{\mathsf{z}}\rangle & \langle B'_{\mathsf{z}}|\tilde{A}_{\mathsf{x}}\rangle\\ \langle B'_{\mathsf{x}}|\tilde{A}_{\mathsf{z}}\rangle & \langle B'_{\mathsf{x}}|\tilde{A}_{\mathsf{x}}\rangle\end{bmatrix}$$

of $R'=O_BR$ such that $s'_1\ge s'_2$. Then, by using Corollary 11, we have

$$\begin{aligned}&\min_{\tilde{\mathcal{E}}\in\mathcal{P}_{ba}(\omega')}\left[H_{\tilde{\mathcal{E}}}(X|E)-H_{\tilde{\mathcal{E}}}(X|Y)\right]\\&\le G_1\left(\mathcal{E}'\right)\\&=1-h\left(\frac{1+s'_1}{2}\right)-h\left(\frac{1+s'_2}{2}\right).\end{aligned}\tag{F.1}$$

By using the minimax principle for singular values [3, Problem 3.6.1], we have

$$\begin{aligned}s'_1&=\max_{x\in\mathbb{R}^2:\|x\|=1}\|S'x\|\\&=\max_{\substack{\alpha,\beta\in\mathbb{R}\\ \alpha^2+\beta^2=1}}\left\|\begin{bmatrix}\langle B'_{\mathsf{z}}|\tilde{A}_{\mathsf{z}}\rangle & \langle B'_{\mathsf{z}}|\tilde{A}_{\mathsf{x}}\rangle\\ \langle B'_{\mathsf{x}}|\tilde{A}_{\mathsf{z}}\rangle & \langle B'_{\mathsf{x}}|\tilde{A}_{\mathsf{x}}\rangle\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\right\|\\&=\max_{\substack{\alpha,\beta\in\mathbb{R}\\ \alpha^2+\beta^2=1}}\left\|\begin{bmatrix}\langle B'_{\mathsf{z}}|(\alpha|\tilde{A}_{\mathsf{z}}\rangle+\beta|\tilde{A}_{\mathsf{x}}\rangle)\\ \langle B'_{\mathsf{x}}|(\alpha|\tilde{A}_{\mathsf{z}}\rangle+\beta|\tilde{A}_{\mathsf{x}}\rangle)\end{bmatrix}\right\|\\&=\max_{\substack{\alpha,\beta\in\mathbb{R}\\ \alpha^2+\beta^2=1}}\sqrt{\langle B'_{\mathsf{z}}|\Gamma_{\alpha,\beta}\rangle^2+\langle B'_{\mathsf{x}}|\Gamma_{\alpha,\beta}\rangle^2}\\&\le\max_{\substack{\alpha,\beta\in\mathbb{R}\\ \alpha^2+\beta^2=1}}\sqrt{\langle B^*_{\mathsf{z}}|\Gamma_{\alpha,\beta}\rangle^2+\langle B^*_{\mathsf{x}}|\Gamma_{\alpha,\beta}\rangle^2}\end{aligned}\tag{F.2}$$

$$=s^*_1,\tag{F.3}$$

where we set $|\Gamma_{\alpha,\beta}\rangle := \alpha|\tilde{A}_z\rangle + \beta|\tilde{A}_x\rangle$, and the equality in Eq. (F.2) holds if $|B'_z\rangle$ and $|B'_x\rangle$ span the same subspace as that spanned by $|\tilde{A}_z\rangle$ and $|\tilde{A}_x\rangle$. By using the minimax principle for singular values in a similar manner, we also have

$$s'_2 = \min_{x\in\mathbb{R}^2:\|x\|=1} \|S'x\| \leq s^*_2. \tag{F.4}$$

Combining Eqs. (F.1), (F.3), and (F.4), we have shown that Eq. (34) is smaller than or equals to Eq. (36). □

References

[1] C.H. Bennett, G. Brassard, Quantum cryptography: public key distribution and coin tossing, in: Proceedings of IEEE International Conference on Computers Systems and Signal Processing, Bangalore, India, 1984, pp. 175–179.
[2] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, W.K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A 54 (1996) 3824–3851, http://dx.doi.org/10.1103/PhysRevA.54.3824.
[3] R. Bhatia, Matrix Analysis, Graduate Texts in Mathematics, vol. 169, Springer, 1997.
[4] J.C. Boileau, D. Gottesman, R. Laflamme, D. Poulin, R.W. Spekkens, Robust polarization-based quantum key distribution over a collective-noise channel, Phys. Rev. Lett. 92 (2004) 017901, http://dx.doi.org/10.1103/PhysRevLett.92.017901.
[5] D. Bruß, Optimal eavesdropping in quantum cryptography with six states, Phys. Rev. Lett. 81 (1998) 3018–3021, http://dx.doi.org/10.1103/PhysRevLett.81.3018.
[6] J. Chen, G. Wu, X. Gu, H. Zeng, Stable quantum key distribution with active polarization control based on time-division multiplexing, New J. Phys. 11 (2009) 065004, http://dx.doi.org/10.1088/1367-2630/11/6/065004.
[7] J. Chen, G. Wu, Y. Li, E. Wu, H. Zeng, Active polarization stabilization in optical fibers suitable for quantum key distribution, Opt. Express 15 (2007) 17928–17936, http://dx.doi.org/10.1364/OE.15.017928.
[8] I.L. Chuang, M.A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, J. Modern Opt. 44 (1997) 2455–2467, http://dx.doi.org/10.1080/09500349708231894.
[9] T.M. Cover, J.A. Thomas, Elements of Information Theory, 2nd ed., John Wiley & Sons, 2006.
[10] J.D. Franson, B.C. Jacobs, Operational system for quantum cryptography, Electron. Lett. 31 (1995) 232–234, http://dx.doi.org/10.1049/el:19950153.
[11] A. Fujiwara, P. Algoet, One-to-one parametrization of quantum channels, Phys. Rev. A 59 (1999) 3290–3294, http://dx.doi.org/10.1103/PhysRevA.59.3290.
[12] A. Fujiwara, H. Nagaoka, Operational capacity and pseudoclassicality of a quantum channel, IEEE Trans. Inform. Theory 44 (1998) 1071–1086, http://dx.doi.org/10.1109/18.669165.
[13] N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, Quantum cryptography, Rev. Modern Phys. 74 (2002) 145–195, http://dx.doi.org/10.1103/RevModPhys.74.145.
[14] D. Gottesman, H.K. Lo, Proof of security of quantum key distribution with two-way classical communication, IEEE Trans. Inform. Theory 49 (2003) 457–475, http://dx.doi.org/10.1109/TIT.2002.807289.
[15] B. Kraus, N. Gisin, R. Renner, Lower and upper bounds on the secret-key rate for quantum key distribution protocols using one-way classical communication, Phys. Rev. Lett. 95 (2005) 080501, http://dx.doi.org/10.1103/PhysRevLett.95.080501.
[16] I. Lucio-Martinez, P. Chan, X. Mo, S. Hosier, W. Tittel, Proof-of-concept of real-world quantum key distribution with quantum frames, New J. Phys. 11 (2009) 095001, http://dx.doi.org/10.1088/1367-2630/11/9/095001.
[17] A. Laing, V. Scarani, J.G. Rarity, J.L. O'Brien, Reference-frame-independent quantum key distribution, Phys. Rev. A 82 (2010) 012304, http://dx.doi.org/10.1103/PhysRevA.82.012304.
[18] L. Ma, H. Xu, X. Tang, Polarization recovery and auto-compensation in quantum key distribution network, Proc. SPIE 6305 (2006) 630513, http://dx.doi.org/10.1117/12.679575.
[19] M.A. Nielsen, I.L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, 2000.
[20] J.F. Poyatos, J.I. Cirac, P. Zoller, Complete characterization of a quantum process: the two-bit quantum gate, Phys. Rev. Lett. 78 (1997) 390–393, http://dx.doi.org/10.1103/PhysRevLett.78.390.
[21] R. Renner, Security of quantum key distribution, Ph.D. thesis, Dipl. Phys., ETH, Switzerland, 2005.
[22] R. Renner, N. Gisin, B. Kraus, Information-theoretic security proof for quantum-key-distribution protocols, Phys. Rev. A 72 (2005) 012332, http://dx.doi.org/10.1103/PhysRevA.72.012332.
[23] R.C. Thompson, Principal submatrices IX: interlacing inequalities for singular values of submatrices, Linear Algebra Appl. 5 (1972) 1–12.
[24] A. Trifonov, A. Zavriyev, D. Subacius, Active stabilization of a one-way QKD system, 2009. US patent: Document No. 7587049.
[25] Z.D. Walton, A.F. Abouraddy, A.V. Sergienko, B.E.A. Saleh, M.C. Teich, Decoherence-free subspace in quantum key distribution, Phys. Rev. Lett. 91 (2003) 087901, http://dx.doi.org/10.1103/PhysRevLett.91.087901.
[26] S. Watanabe, R. Matsumoto, T. Uyematsu, Tomography increases key rates of quantum-key-distribution protocols, Phys. Rev. A 78 (2008) 042316, http://dx.doi.org/10.1103/PhysRevA.78.042316.
[27] S. Watanabe, R. Matsumoto, T. Uyematsu, Y. Kawano, Key rate of quantum key distribution with hashed two-way classical communication, Phys. Rev. A 76 (2007) 032312, http://dx.doi.org/10.1103/PhysRevA.76.032312.
[28] L. Wei-Tao, W. Wei, L. Lin-Mei, L. Cheng-Zu, Polarization encoded quantum key distribution over special optical fibres, Chinese Phys. Lett. 23 (2006) 287–289, http://dx.doi.org/10.1088/0256-307X/23/2/004.
[29] G.B. Xavier, V. de Faria, G.P. Temporao, J.P. von der Weid, Full polarization control for fiber optical quantum communication systems using polarization encoding, Opt. Express 16 (2008) 1867–1873, http://dx.doi.org/10.1364/OE.16.001867.
[30] G.B. Xavier, N. Walenta, G.V. de Faria, G.P. Temporão, N. Gisin, Z. Zbinden, J.P. von der Weid, Experimental polarization encoded quantum key distribution over optical fibres with real-time continuous birefringence compensation, New J. Phys. 11 (2009) 045015, http://dx.doi.org/10.1088/1367-2630/11/4/045015.
[31] A. Zavriyev, A. Trifonov, M. Lagasse, Two-way QKD system with active compensation, 2009. US patent: Document No. 7606371.